

**Claims**

1. Method for controlling a speech dialog system, wherein an acoustic output signal is provided in response to an acoustic input signal, comprising the steps of:  
  
receiving a further acoustic input signal,  
  
processing the further acoustic input signal by a voice activity detector to detect voice activity,  
  
processing the further acoustic input signal or an output signal corresponding to the further acoustic input signal provided by the voice activity detector by a speech recognition unit to detect speech, if voice activity was detected by the voice activity detector,  
  
modifying the acoustic output signal if speech was detected by the speech recognition unit during the output of the output signal.
2. Method according to claim 1, wherein the modifying step comprises reducing the volume of the output signal.
3. Method according to claim 1 or 2, wherein the modifying step comprises interrupting the outputting of the output signal.
4. Method according to one of the preceding claims, wherein in the processing step by the speech recognition unit, speech is detected using at least one criterion based on a Hidden Markov Model, a pause model, an artificial neural network, confidence features, the number of interrupted words, and/or a code book.
5. Method according to claim 4, wherein at least two criteria are used, further comprising the step of feeding the results of the at least two criteria to a classification unit.

6. Method according to one of the preceding claims, wherein the processing step by the speech recognition unit comprises recognizing speech by the speech recognition unit.
7. Method according to one of the preceding claims, wherein the receiving step comprises processing the further acoustic input signal by an acoustic echo canceller, a noise reduction means, and/or a feedback suppression means.
8. Method according to one of the preceding claims, wherein the receiving step comprises receiving a plurality of further acoustic input signal emanating from a plurality of microphones.
9. Method according to claim 7, wherein the receiving step comprises combining the plurality of input signals, preferably using a beamformer.
10. Computer program product directly loadable into an internal memory of a digital computer, comprising software code portions for performing the steps of the method according to one of the claims 1 to 9.
11. Computer program product stored on a medium readable by a computer system, comprising computer readable program means for causing a computer to perform the steps of the method according to one of the claims 1 to 9.
12. Speech dialog system, comprising a signal input unit, a voice activity detector, a speech recognition unit, and a signal output unit,  
  
wherein the speech dialog system is configured such that if the voice activity detector detects voice activity for an input signal and the speech recognition unit detects speech for the input signal or for an output signal corresponding to the input signal provided by the voice activity detector during

an output of the signal output unit, the output signal of the signal output unit is modified.

13. Speech dialog system according to claim 12, wherein the output signal comprises a speech signal.
14. Speech dialog system according to claim 12 or 13, wherein the speech dialog system is configured such that the modification of the output signal is a reduction of the volume and/or an interruption of the output signal.
15. Speech dialog system according to one of the claims 12 – 14, further comprising a control unit, wherein  
  
the voice activity detector has an output for providing a detector output signal if speech activity was detected,  
  
the speech recognition unit has an input for receiving the detector output signal and an output for providing a recognizer output signal if speech was detected,  
  
the control unit has an input for receiving the recognizer output signal and an output for providing a control signal depending on the recognizer output signal, and  
  
the signal output unit has an input for receiving the control signal and an output for providing an output signal depending on the control signal.
16. Speech dialog system according to claim 15, wherein the control signal initiates an output of an output signal.
17. Speech dialog system according to claim 15 or 16, wherein

the speech recognition unit is configured to determine a modification signal if speech is detected by the speech recognition unit and further comprises an output for providing the modification signal, and

the signal output unit is connected to the speech recognition unit and comprises an input for receiving the modification signal.

18. Speech dialog system according to claim 17, wherein the modification signal is configured to interrupt the output of an output signal.
19. Speech dialog system according to one of the claims 12 – 18, wherein the signal input unit comprises a plurality of microphones and a beamformer and comprises an output for providing a beamformed input signal to the signal detection unit.
20. Speech dialog system according to one of the claims 12 – 19, wherein the signal input unit comprises echo cancellation means and/or noise reduction means and/or feedback suppression means.
21. Speech dialog system according to one of the claims 12 – 20, wherein the signal output unit further comprises a memory for storing at least one predetermined output signal and/or a signal synthesizing means, preferably a speech synthesizing means.